УДК 378:004

СИСТЕМА ГЕНЕРАЦІЇ СИГНАТУР ВИЯВЛЕННЯ ШКІДЛИВОГО ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

М. О. Женжера, магістрант
Київський національний університет технологій та дизайну
Т.І. Демківська, кандидат технічних наук, доцент
Київський національний університет технологій та дизайну

Ключові слова: шкідливе програмне забезпечення, антивірусні програми, виявлення загроз, сигнатурне виявлення, статичний аналіз, динамічний аналіз.

Основною метою дослідницького проекту є розроблення проєкту для генерації сигнатур виявлення шкідливого програмного забезпечення.

Проєкт складається з наступних основних компонентів:

- (**enrichment**) збагачення бази даних з різних джерел;
- (**generation**) генерація сигнатур на основі підготовленої бази даних;
- (**export**) експорт якісних згенерованих сигнатур у потрібному форматі.

Важливим принципом в розробленні цього проєкту є універсальність, гнучкість. Джерела, формати та додаткові правила повинні мати можливість легкого налаштування та заміни між собою.

Розглянемо кожен компонент більш детально:

1. (**enrichment**) збагачення бази даних з різних джерел;

Основною моделлю проєкта є шкідливий виконуваний файл, а також інші артефакти (мережева активність, ключі реєстру, тощо).

Головним джерелом даних для цього проєкту є сервіси динамічного аналізу програмного забезпечення з використанням віртуальних машин (**sandbox**). Також проєкт має мати можливість розширення іншими джерелами як, наприклад, пропрієтарні дані (логи) та звіти в інших форматах.

2. (**generation**) генерація сигнатур на основі підготовленої бази даних;

Ключовий компонент цього проєкта. Мета цього компонента - отримати на виході сигнатури, що готові для використання, отримуючи дані з різних джерел на вході. Сигнатури мають бути достатньо перевіреними, що виключить ризик масованого помилкового виявлення доброякісних файлів.

3. (**export**) експорт якісних згенерованих сигнатур у потрібному форматі.

Для того щоб скористатися результатом, його потрібно надати в якомусь форматі. А окрім формату, важливим критерієм цього компоненту є можливість пошуку та експорту лише потрібних даних.

Щоб ідентифікувати шкідливе програмне забезпечення, майже кожен сканер шкідливого програмного забезпечення використовує певну комбінацію методів на основі сигнатур і аномалій. Однак існують способи,

які використовують переваги цих методів, зокрема статичні методи, які виконуються шляхом вилучення характеристик із статичного шкідливого програмного забезпечення, яке зберігається на диску, і гібридні методи, які поєднують динамічний і статичний підходи.

Щоб визначити, чи є файл шкідливим, процедури на основі сигнатур використовують скануюче програмне забезпечення для порівняння вмісту файлу з базою даних, що містить тисячі вірусних сигнатур. Основною перевагою цих методів є їхня ефективність.

Генерація сигнатур, як ключовий етап, також повинна бути адаптивною. Це передбачає можливість використання різних алгоритмів залежно від характеру загрози або типу даних, що обробляються. Наприклад, для аналізу шкідливого ПЗ, що має складні методи захисту, можуть використовуватися алгоритми на основі поведінкових моделей, тоді як для простих загроз – класичні статичні підходи. Така гнучкість дозволить системі адаптуватися під різні умови експлуатації.

Окрему увагу варто приділити процесу експорту сигнатур, оскільки від цього залежить подальша інтеграція з іншими системами. Для забезпечення універсальності проекту необхідно реалізувати підтримку різних форматів вивантаження даних – як стандартних форматів (наприклад, JSON або XML), так і специфічних для певного обладнання або програмного забезпечення. Окрім цього, важливо, щоб система могла працювати в режимі реального часу, надаючи згенеровані сигнатури відразу після їхнього створення. Це дозволить мінімізувати час від виявлення загрози до її нейтралізації.

Також у проекті передбачено можливість додавання нових правил та джерел даних без значного втручання у базову архітектуру системи. Це сприяє її універсальності та довготривалому використанню, оскільки нові методи виявлення загроз можуть інтегруватися без необхідності повної перебудови всієї системи. Така модульність значно спрощує підтримку та розвиток проекту, що робить його більш гнучким та стійким до змін у галузі кібербезпеки.

Таким чином, запропонований проект не тільки покращує традиційні підходи до виявлення загроз, але й створює гнучку платформу, здатну адаптуватися до швидко змінюваного ландшафту кіберзагроз. Застосування різних методів та джерел даних дозволить забезпечити більш комплексний підхід до виявлення шкідливого ПЗ, що у свою чергу підвищить надійність системи безпеки.

Список використаних джерел

1. Abiola, Alogba & Marhusin, M.F. (2018). Signature-Based Malware Detection Using Sequences of N-grams. International Journal of Engineering and Technology(UAE). 7. 10.14419/ijet.v7i4.15.21432.

2. Agrawal, Rahul & sharma, yogesh & Awasthi, Harsh. (2021). A SIGNATURE-BASED MALWARE DETECTION SYSTEM. 2. 115.

3. Lockett, Adam. (2021). Assessing the Effectiveness of YARA Rules for Signature-Based Malware Detection and Classification. 10.48550/arXiv.2111.13910.